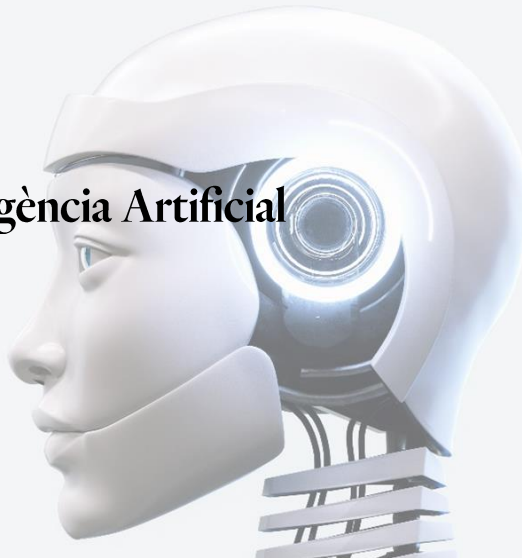


RocaJunyent Informa

Departament de Protecció de Dades, Ciberseguretat i Tecnologia

19 de juliol de 2024

El nou Reglament d'Intel·ligència Artificial



El passat 12 de juliol de 2024 es va publicar el text definitiu del nou Reglament d'Intel·ligència Artificial ("RIA" o "Reglament") en el Diari Oficial de la Unió Europea. Analitzem els seus punts clau en aquesta alerta.

1. Què considera el RIA com a Sistemes d'Intel·ligència Artificial?

El RIA està dissenyat per a supervisar i regular la implementació de la Intel·ligència Artificial ("IA"), enfocant-se particularment a mitigar els riscos associats amb el seu desenvolupament i ús.

El Reglament defineix els **Sistemes de IA** com a sistemes basats en una màquina que està dissenyada per a funcionar amb diferents nivells d'autonomia i que pot mostrar capacitat d'adaptació després del desplegament, i que, per a objectius explícits o implícits, infereix de la informació d'entrada que rep la

manera de generar resultats de sortida, com a prediccions, continguts, recomanacions o decisions, que poden influir en entorns físics o virtuals.

Així mateix, el RIA posa l'accent principalment en els **Sistema de IA d'Ús General**, definint-los com a sistemes basats en un **Model de IA d'Ús General**, que està entrenat amb un gran volum de

dades utilitzant autosupervisió a gran escala, que presenta un grau considerable de generalitat i és capaç de realitzar de manera competent una gran varietat de tasques distintes, independentment de la manera en què el model s'introdueixi en el mercat, i que pot integrar-se en diversos sistemes o aplicacions posteriors, excepte els models de IA que s'utilitzen per a activitats de recerca, desenvolupament o creació de prototips abans de la seva introducció en el mercat; els Sistemes de IA d'Ús General pot servir per a diversos fins, tant per al seu ús directe com per a la seva integració en altres sistemes de IA.

2. Quins rols identifica el RIA?

El RIA identifica els següents rols en el cicle de vida dels sistemes i models IA:

- a) **Proveïdor:** Persona física o jurídica, autoritat pública o organisme que desenvolupa un sistema o model de IA i ho introdueix en el mercat o posa en servei sota el seu propi nom o marca, ja sigui de manera gratuïta o pagada.
- b) **Responsable de desplegament:** Persona física o jurídica, o autoritat pública que utilitza un sistema de IA sota la seva pròpia autoritat, excepte quan l'ús és de caràcter personal i no professional.
- c) **Importador:** Persona física o jurídica situada a la Unió Europea que introdueix en el mercat un sistema de IA que porta el nom o la marca d'una entitat establerta fora de la UE.
- d) **Distribuïdor:** Persona física o jurídica en la cadena de subministrament, distinta del proveïdor o importador, que comercialitza un sistema de IA a la Unió Europea.
- e) **Representant autoritzat:** Persona física o jurídica a la Unió Europea que, mitjançant mandat escrit d'un proveïdor, assumeix les obligacions i procediments establerts en el reglament en representació del proveïdor.
- f) **Fabricant del producte:** Fabricant segons la definició de la legislació harmonitzada llistada en l'Annex I del RIA.

Els sis rols es denominen conjuntament “operador” i tots ells poden aplicar-se al context dels sistemes de IA. No obstant això, en el context dels models, només es poden aplicar els rols de proveïdor o representant autoritzat. És a dir, no és possible ser responsable de desplegament, importador, distribuïdor o fabricant del producte respecte a un model de IA

3. Quin és l'àmbit d'aplicació?

L'àmbit d'aplicació del Reglament s'estén més enllà de les organitzacions establertes a la Unió Europea (“UE”), aconseguint també a aquelles que utilitzin, desenvolupin o intervinguin com a operadors en els sistemes i models d'Intel·ligència Artificial (“IA”) en el mercat europeu. Això implica que, també hauran d'aplicar-ho les nombroses multinacionals, amb presència o interaccions comercials a la UE.

4. A quins sistemes IA no s'aplica el RIA?

El Reglament de IA no s'aplicarà als sistemes i eines de IA de dedicats a:

- a) **Fins militars i de seguretat nacional:** Sistemes de IA utilitzats exclusivament per a fins militars, de defensa o de seguretat nacional.
- b) **Cooperació internacional:** Sistemes de IA utilitzats per autoritats públiques de tercers països o organitzacions internacionals en el marc d'acords de cooperació internacional amb la UE o els seus Estats membres, sempre que es garanteixi la protecció dels drets i llibertats fonamentals.

- c) **Recerca i desenvolupament científic:** Sistemes i models de IA desenvolupats i utilitzats exclusivament per a la recerca i el desenvolupament científics.
- d) **Proves i desenvolupament precomercial:** Activitats de recerca, prova o desenvolupament de sistemes o models de IA abans de la seva introducció en el mercat o posada en servei (excepte proves en condicions reals).
- e) **Ús personal no professional:** Sistemes de IA utilitzats per persones físiques en activitats purament personals i no professionals.
- f) **Llicències lliures i codi obert:** Alguns sistemes de IA alliberats baix llicències lliures i de codi obert, excepte aquells que siguin prohibits o d'alt risc.

5. Com es categoritzen els Sistemes IA?

Els Sistemes IA es categoritzen segons el risc que impliquen per als drets fonamentals de les persones, existint Sistemes de Risc Inacceptable, d'Alt Risc, de Risc Limitat i de Risc Mínim.

Risc Inacceptable:

Els següents Sistemes IA són catalogats de Risc Inacceptable i, per tant, estan prohibits pel RIA:

1. **Inferència de característiques sensibles:** La categorització biomètrica i individual per a inferir característiques sensibles com la raça, orientació religiosa, ideològica o sexual, entre altres, en plantejar seriosos riscos per a la privacitat i la discriminació.
2. **Reconeixement biomètric en espais públics:** L'ús de reconeixement biomètric en remot i en temps real en espais accessibles al públic amb finalitats de prevenció o recerca de delictes, sense complir uns certs requisits d'excepció (com a tipus penals específics i condicions estrictes).
3. **Scraping d'imatges facials:** El scraping generalitzat d'imatges facials d'internet o CCTV per a generar o ampliar bases de dades de reconeixement facial.
4. **Reconeixement d'emocions en entorns sensibles:** El reconeixement d'emocions en entorns laborals i educatius, salvo amb finalitats mèdiques o de seguretat,
5. **Scoring social:** Els sistemes de puntuació social que avaluen a persones o col·lectius segons el seu comportament social o característiques personals (conegudes, inferides o predites).
6. **Manipulació:** Els sistemes que utilitzen tècniques subliminals, manipuladores o enganyoses per a alterar el comportament de les persones, minvant la seva capacitat per a prendre decisions informades i portant-les a adoptar decisions perjudicials que d'una altra manera no haguessin pres.
7. **Explotació de vulnerabilitats:** Sistemes que explotin vulnerabilitats de les persones (edat, solvència, discapacitat, etc.) o grups de persones per a influir en el seu comportament provocant perjudicis considerables.
8. **Predicció de risc delictiu:** Sistemes que avaluin i prediguin el risc de comissió de delictes basant-se únicament en l'elaboració del perfil d'una persona física.

Alt Risc

Un sistema de IA és considerat d'alt risc si compleix alguna de les següents condicions:

1. Està destinat a ser un component de seguretat d'un producte regulat per la legislació harmonitzada de la UE (Annex I), o és el propi producte.
2. El producte o el sistema de IA ha de sotmetre's a una avaluació de conformitat de tercers segons la legislació harmonitzada de la UE (Annex I).
3. Pertany a algun dels vuit sectors identificats en l'Annex III del RIA i poden presentar un risc significatiu per a la salut, seguretat o drets fonamentals, ni influeixen substancialment en la presa de decisions
 - a. Biometria
 - b. Infraestructures Críques
 - c. Educació i Formació professional
 - d. Ocupació, gestió dels treballadors i accés a l'autoocupació
 - e. Accés a serveis privats essencials i a serveis i prestacions públics essencials i gaudi d'aquests serveis i prestacions
 - f. Garantia del compliment del dret
 - g. Migració, asil i gestió del control fronterer
 - h. Administració de justícia i processos democràtics

Els Proveïdors que considerin que un sistema pertanyent als sectors descrits no és d'alt risc ha de documentar la seva avaluació abans de la seva comercialització i registrar-la

6. Com es categoritzen els Models de IA ús general?

a) Risc Sistèmic:

Un model de IA d'ús general es considera de risc sistèmic si compleix alguna de les següents condicions:

1. Capacitats de gran impacte avaluades mitjançant eines i metodologies tècniques adequades, utilitzant indicadors i paràmetres de referència.
2. Decisió de la Comissió Europea en atenció a criteris taxats:
 - a. El nombre de paràmetres del model;
 - b. La qualitat o la grandària del conjunt de dades, per exemple mesurats a través de criptofichas;
 - c. La quantitat de càlcul utilitzada per a entrenar el model, mesura en operacions de coma flotant o indicada amb una combinació d'altres variables, com el cost estimat de l'entrenament, el temps estimat necessari o el consum d'energia estimat per a aquest;
 - d. Les modalitats d'entrada i sortida del model, com la conversió de text a text (grans models de llenguatge), la conversió de text a imatge, la multimodalitat i els l·lindars capdavaners per a determinar les capacitats de gran impacte de cada modalitat, i el tipus concret d'entrades i sortides (per exemple, seqüències biològiques);

- e. Els paràmetres de referència i les avaluacions de les capacitats del model, també tenint en compte el nombre de tasques sense entrenament addicional, l'adaptabilitat per a aprendre tasques noves distintes, el seu nivell d'autonomia i capacitat d'ampliació i les eines a les quals té accés;
- f. Si les seves repercussions per al mercat interior són importants a causa del seu abast, la qual cosa es donarà per descomptat quan s'hagi posat a la disposició d'almenys 10 000 usuaris professionals registrats establerts a la UE;
- g. El nombre d'usuaris finals registrats.

b) Riscos Ordinaris

Tots els models no qualificats com de risc sistèmic.

7. Què són els espais de prova i com funcionaran?

Amb l'objectiu de reforçar la competitivitat europea en l'àmbit de la IA, s'habilita un espai controlat de proves ("Sandbox") orientat a oferir un entorn controlat per a desenvolupar, provar, i validar sistemes de IA innovadors en condicions reals. Amb el Reglament, es busca establir normes comunes per a la creació d'un Sandbox i un marc per a la cooperació entre les autoritats de supervisió implicades. No obstant això, seran les autoritats de supervisió locals els qui determinaran els requisits concrets per a l'accés als Sandbox. A Espanya, el Reial decret 817/2023, de 8 de novembre, té com a objecte la creació del primer entorn controlat de per a assajar l'aplicació d'uns certs requisits previstos en el Reglament

8. Què són els codis de conducta que s'esmenten en el Reglament de IA?

Els codis de conducta són eines per a promoure l'ètica, la transparència i la responsabilitat en el desenvolupament i ús de la IA. L'Oficina de IA i els Estats membres fomentaran i facilitaran l'elaboració de codis de conducta, dissenyats per a ser aplicats de manera voluntària pels proveïdors de sistemes de IA. Aquests codis estan destinats a fomentar l'aplicació voluntària dels requisits establerts en el capítol III, secció 2, als sistemes de IA que no siguin d'alt risc. Aquests hauran d'incloure, entre altres, mecanismes de governança, les millors pràctiques del sector i l'adopció de solucions tècniques.

9. Quan serà aplicable?

El RIA serà aplicable de manera gradual:

- 1r) 1 d'agost de 2024:** Entrada en vigor del RIA
- 2n) 2 de febrer de 2025:** Entrada en aplicació de les prohibicions d'ús, sis mesos després de l'entrada en vigor del reglament.
- 3r) 2 de maig de 2025:** Conclou el termini per a la finalització dels codis de bones pràctiques, nou mesos després de l'entrada en vigor del reglament.
- 4t) 2 d'agost de 2025:** Entrada en aplicació de les disposicions dirigides a models de IA, dotze mesos després de l'entrada en vigor del reglament.
- 5è) 2 d'agost de 2026:** Entrada en aplicació de la resta de la norma, vint-i-quatre mesos després de l'entrada en vigor del reglament.
- 6è) 2 d'agost de 2027:** Entrada en aplicació de l'article 6.1 i de l'annex I del RIA per a qualificar sistemes d'Alt Risc, trenta-sis mesos després de l'entrada en vigor del reglament.

Als 5 anys s'haurà de realitzar una revisió del Reglament de IA.

10. Quines són les entitats de supervisió?

A Espanya, l'autoritat de supervisió competent per al Reglament d'Intel·ligència Artificial és la AESIA (Agència Espanyola de Supervisió de la Intel·ligència Artificial). A nivell europeu, s'establirà un Comitè Europeu d'Intel·ligència Artificial amb un representant de cada Estat membre, encarregat d'orientar la implementació del reglament, elaborar guies i establir regles per a sandboxes. L'Oficina Europea de IA exerceix un paper crucial a supervisar el compliment del reglament en els Estats membres, assegurant una aplicació uniforme i efectiva en tots els sectors que utilitzen sistemes de IA. Treballa en col·laboració amb les autoritats nacionals per a coordinar activitats de control i formació, garantint el respecte als estàndards establerts. A nivell internacional, l'Oficina participa en fòrums i acords globals per a promoure una governança global de IA i compartir estratègies reguladores comunes.

11. Quins són els incompliments i les sancions?

Les multes per infracció del RIA es van fixar com un percentatge del volum de negocis anual global de l'empresa infractora en l'exercici financer anterior o una quantitat predeterminada, la que fora més alta. Com a màxim representaria el 7% del volum de negoci global. Les mesures punitives estan dissenyades per a garantir el compliment del RIA i fomentar un ús ètic i responsable de la intel·ligència artificial. Algunes de les sancions que podrien aplicar-se són les següents:

En Models de IA, fins a 15M d'euros o el 3% pel volum de negoci per a l'incompliment de les obligacions

En els Sistemes de IA:

- Fins a 35M de €, o 7% del volum anual de negoci per a empreses per incomplir les prohibicions del Reglament.
- Fins a 15M de €, o 3% del volum anual de negoci per incomplir les obligacions del Reglament a proveïdors, importadors, distribuïdors, usuaris.
- Fins a 7,5M de €, o 1% del volum anual de negoci per subministrar informació incorrecta a entitats notificades o autoritats nacionals competents

A més de les sancions econòmiques, les empreses poden enfrontar-se a una sèrie de conseqüències derivades de les infraccions del RIA. Prohibició d'ús. En casos greus d'incompliment, les autoritats competents poden imposar la prohibició de l'ús d'uns certs sistemes d'intel·ligència artificial o de tota l'activitat relacionada amb la IA a l'empresa infractora. Retirada de certificacions i llicències. Si una empresa ha obtingut certificacions o llicències per a l'ús de sistemes d'intel·ligència artificial i es descobreix que ha incomplert les regulacions del RIA, aquestes certificacions o llicències poden ser revocades. Responsabilitat civil i penal. A més de les sancions administratives, les empreses i persones responsables del desenvolupament o ús indegut de sistemes de IA poden enfrontar-se a accions legals civils i penals per danys causats a tercers o per violacions de drets fonamentals

ⁱ [Reglamento \(UE\) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos \(CE\) n.º 300/2008, \(UE\) n.º 167/2013, \(UE\) n.º 168/2013, \(UE\) 2018/858, \(UE\) 2018/1139 y \(UE\) 2019/2144 y las Directivas 2014/90/UE, \(UE\) 2016/797 y \(UE\) 2020/1828 \(Reglamento de Inteligencia Artificial\)](#)